



After

PTO/SB/21 (09-04)

Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM (to be used for all correspondence after initial filing)	Application Number	09/275,568
	Filing Date	3/24/1999
	First Named Inventor	Michael C. Pitman
	Art Unit	2168
	Examiner Name	Cheyne D. Ly
Total Number of Pages in This Submission	Attorney Docket Number	YOR919980112US1

ENCLOSURES (Check all that apply)		
<input type="checkbox"/> Fee Transmittal Form <input type="checkbox"/> Fee Attached <input type="checkbox"/> Amendment/Reply <input type="checkbox"/> After Final <input type="checkbox"/> Affidavits/declaration(s) <input type="checkbox"/> Extension of Time Request <input type="checkbox"/> Express Abandonment Request <input type="checkbox"/> Information Disclosure Statement <input type="checkbox"/> Certified Copy of Priority Document(s) <input type="checkbox"/> Reply to Missing Parts/ Incomplete Application <input type="checkbox"/> Reply to Missing Parts under 37 CFR 1.52 or 1.53	<input type="checkbox"/> Drawing(s) <input type="checkbox"/> Licensing-related Papers <input type="checkbox"/> Petition <input type="checkbox"/> Petition to Convert to a Provisional Application <input type="checkbox"/> Power of Attorney, Revocation <input type="checkbox"/> Change of Correspondence Address <input type="checkbox"/> Terminal Disclaimer <input type="checkbox"/> Request for Refund <input type="checkbox"/> CD, Number of CD(s) _____ <input type="checkbox"/> Landscape Table on CD	<input type="checkbox"/> After Allowance Communication to TC <input type="checkbox"/> Appeal Communication to Board of Appeals and Interferences <input checked="" type="checkbox"/> Appeal Communication to TC (Appeal Notice, Brief, Reply Brief) <input type="checkbox"/> Proprietary Information <input type="checkbox"/> Status Letter <input checked="" type="checkbox"/> Other Enclosure(s) (please identify below): Postcard
<div>Remarks</div>		

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT			
Firm Name	Holland & Knight LLP		
Signature	<i>Michael J. Buchenhorner</i>		
Printed name	Michael J. Buchenhorner		
Date	<i>July 17, 2006</i>	Reg. No.	33,162

CERTIFICATE OF TRANSMISSION/MAILING			
I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below:			
Signature	<i>Michael J. Buchenhorner</i>		
Typed or printed name	Michael J. Buchenhorner	Date	<i>July 17, 2006</i>

This collection of information is required by 37 CFR 1.5. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450,

On July 17, 2006

Date of Deposit

Michael J. Buchenhour
Signature

In re Application of: PITMAN et al.

Serial No.: 09/275,568
Filed: 03/24/99

Group Art Unit: 1631
Examiner: Cheyne D. Ly

Title: Similarity Searching of Molecules Based Upon Descriptor Vectors Characterizing Molecular Regions

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

SECOND AMENDED APPEAL BRIEF

Sir:

This Appeal Brief is submitted in response to a Notification of Non-Compliant Appeal Brief dated June 16, 2006. The Notification required a correct statement of the status of all claims; a statement that at least one amendment has been filed after the final rejection; a corrected copy of the appealed claims; and other reasons stated in a notification of non-compliant brief in a prior appeal.

REAL PARTY IN INTEREST

International Business Machines Corporation is the real party in interest as assignee of the subject application.

RELATED APPEALS AND INTERFERENCES

Appellant has previously appealed a prior final action. No decision was made by the Board because the Examiner issued a non-final action after a second amended appeal brief was filed.

STATUS OF CLAIMS

Claims 1 and 4-15, and 31-35 (the claims at issue) are pending in the above-identified patent application and their rejection is appealed. Claims 2-3 were cancelled in an amendment dated November 8, 2002. Claims 16-30 were cancelled in an amendment dated November 2, 2001. Claims 31-35 were added in an amendment filed September 15, 2005 and rejected in the final office action. The claims at issue were finally rejected in an Office Action dated June 29, 2005. The final rejection of the claims at issue is hereby appealed.

STATUS OF AMENDMENTS

On September 1, 2005 Appellant submitted a proposed amendment after final rejection. That amendment was not entered by the Examiner. Appellant is filing a proposed amendment canceling claims 34 and 35 concurrently herewith.

SUMMARY OF INVENTION

Claim 1

The present invention, as set forth in claim 1, and as described and shown in the specification and the Figures of the above-identified patent application, is directed to a method for generating and storing data characterizing at least one region of said plurality of regions. The method comprises the steps of:

generating an entry [page 3, line 18] comprising i) an identifier that identifies said at least one region, and ii) data characterizing a set of axes derived from a property distribution of said at least one region [page 3, lines 19-20];

applying a mapping to the descriptor vector associated with said at least one region [page 3, lines 20-21] based on preselected criteria [page 3, lines 10-13];

generating a key that corresponds to said mapping of the descriptor vector associated with said at least one region [page 3, lines 20-21]; and

storing said entry in a memory [page 3, line 21], wherein said key is associated with said entry such that the key indexes the entry for retrieval thereof [page 4, lines 2-3]. A concept underlying the claimed invention is the storage of data in groupings that are sensitive to the way a human would search for stored information, thus facilitating retrieval of the stored data in a way that is useful for using the molecules.

Claim 4

Claim 4 recites the method of claim 1 wherein the property distribution of said at least one region is computed from a convolution with a probe function to a property field [page 29, line 15-page 30, line 5].

Claim 6

Claim 6 recites the method of claim 5 wherein the mapping is derived from the steps of: generating first data representing differences between the groups of descriptor vectors [page 49, line 2]; generating second data representing variations within the groups of descriptor vectors [page 49, line 3]; identifying a set of component vectors that maximizes a ratio of variations between groups to the variations within the groups along the component vectors as a discriminant criterion function [page 9, lines 10-22, and page 23, line 15]; generating a criterion function for subsets of the component vectors [page 10, lines 17], wherein the criterion function utilizes the first data and the second data [page 22, lines 21-page 23 line 7]; for each particular subset of component vectors, calculating a probability value for the criterion functions associated with the particular

subset [page 10, lines 20-24 and page 23, lines 8-22]; selecting a probability value from probability values for said subsets of component vectors based upon a predetermined criterion [page 10, lines 20-24]; identifying the subset of component vectors associated with the selected probability value [Id.]; and generating a mapping to a space corresponding to the subset of component vectors associated with the selected probability value [page 15, lines 7-9], and storing the mapping for subsequent processing.

Claim 8

Claim 8 recites the method of claim 7, wherein the criterion function has the general form:

$$J(\hat{w}) = C \left(\frac{\hat{w}^T \mathbf{\Sigma}_b \hat{w}}{\hat{w}^T \mathbf{\Sigma}_w \hat{w}} \right)$$

where \hat{w} is some vector, T indicates a transpose, $\mathbf{\Sigma}_b$ is a first data representing covariance, $\mathbf{\Sigma}_w$ is a second data representing covariance and C is a constant based upon degrees of freedom in $\mathbf{\Sigma}_b$ and $\mathbf{\Sigma}_w$ [page 8, line 19 -page 20, line 14].

Claim 12

Claim 12 recites the method of claim 11, wherein the statistic for a given subset of component vectors has the following form:

$$\Psi_s = C \left(\frac{1}{L_s} \right) \sum J_k$$

where f_k represents the value of the criterion function at a component vector in the given subset,

C is a constant,

L_S represents the number of f_k values in the given subset of component vectors, and the Σ operation sums over the $L_S f_k$ values in the given subset of component vectors [page 20, lines 4-19].

ISSUES

- I. Whether the Examiner erred in rejecting claims 1, 4-15, and 31-35 as being directed to non-statutory subject matter.
- II. Whether the Examiner erred in rejecting claims 1, 4, and 5 under 35 U.S.C. §§ 102(a) and (e)(2) as being anticipated by U.S. Patent Number 5,784,294 (Platt).
- III. Whether the Examiner erred in rejecting claims 1, 4-15, and 31-35 under 35 U.S.C. §112, first paragraph, as failing to comply with the written description requirement.
- IV. Whether the Examiner erred in rejecting claims 8, 9, and 12-14 under 35 U.S.C. §112, second paragraph, as being indefinite.
- V. Whether the Examiner erred in objecting to the amendment of September 14, 2004 under 35 USC 132(a) as introducing new matter.

ARGUMENT

I. Rejections of Claims 1, 4-15, and 31-35 Under 35 U.S.C. § 101

The Examiner erred in rejecting claims 1, 4-15 and 31-35 under 35 U.S.C. §101 on grounds that the claimed invention is allegedly directed to non-statutory subject matter. The analysis of whether a claim is directed to statutory subject matter begins with the language of 35 U.S.C. 101, which reads:

"Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title."

In AT&T Corp. v. Excel Communications, Inc., 172 F.3d 1352, 50 USPQ2d 1447 (Fed. Cir. 1999), the United States Court of Appeals for the Federal Circuit said that the Supreme Court has construed 101 broadly, noting that Congress intended statutory subject matter to "include anything under the sun that is made by man." See Diamond v. Chakrabarty, 447 U.S. 303, 309 (1980) (quoting S. Rep. No. 82-1979, at 5 (1952); H.R. Rep. No. 82-1923, at 6 (1952)); see also Diamond v. Diehr, 450 U.S. 175, 182 (1981). Notwithstanding the broad scope statutory subject matter, the Court has specifically identified three categories of unpatentable subject matter: "laws of nature, natural phenomena, and abstract ideas." See Diehr, 450 U.S. at 185.

In this appeal, all of the claims at issue are method claims which fall within the "process" category of the four enumerated categories of patentable subject matter in 101. The examiner determined that the claims at issue recite a method that "is merely arranging the data based on an algorithm without any practical application."

The subject patent of the application claims are methods or processes for generating and storing data. Moreover these processes are performed by a machine (a data

processing system). The data are expressed and processed as electrical signals operated upon by a processing apparatus. That is a practical application of the invention. The physical alteration that the Examiner believes to be missing is at least in the storing step which reorganizes the data represented by electrical signals into a different data structure and storing that different (altered) data structure into a memory. The claim also states in the storing step that the key is associated with the subject data entry for retrieval thereof.

The Examiner's determination constitutes an error of law. The issue of failure to claim statutory subject matter is one of law that is reviewed de novo. See AT&T Corp. v. Excel Communications, Inc., supra.

The storage of data in a computer memory is by itself a concrete and useful result. Claim 1 which is representative includes four steps that precisely set forth how the subject information is stored in memory. In In re Lowry, 32 F. 3d 1579, 32 USPQ2d 1031 (Fed. Cir. 1994), the Federal Circuit held that claims directed to an invention related to "storage, use, and management of information residing in a memory were entitled to patentable weight." In Lowry, the Board of Patent Appeals and Interferences reversed the statutory subject matter rejections made by the Examiner under 35 U.S.C. §101. The Federal Circuit then went further by holding that claim limitations relating to the storage of information in a computer memory are entitled to patentable weight in distinguishing the prior art. The method claimed in the instant application is similar to the storage of data in Lowry. Storage of information in a computer memory is an important aspect of information technologies because information processing apparatus must read data from computer memory to execute operations on that information. Considering the claims at issue as a whole, as they must be, it becomes clear that the information stored in a memory has a practical purpose beyond the mere storage of the information – retrieval of the stored information based on the key mapped to a descriptor vector. The ability to retrieve information from memory based on various

criteria is perhaps as important as storage of the information.

The technology of search engines which is the subject of numerous patents is concerned with this very concept. Failure to provide patent protection to inventions in the art of data retrieval would violate the constitutional mandate of promoting the progress of science and the useful arts. Therefore, the rejection of the claims at issue for failure to recite statutory subject matter must be reversed.

In the final office action dated January 27, 2003, the Examiner rejected appellants' arguments and reasserted the position that the claims are directed to unpatentable subject matter. The Examiner thus contended that: "An invention where a system merely stores data such as descriptor vectors associated with a plurality of regions of molecules onto a media [sic, medium] is considered to be non-statutory subject matter because the said data is considered to be nonfunctional descriptive material." This rejection is nothing more than an application of the "printed matter" category of non-patentable subject matter that was so clearly discredited and reversed in the In re Lowry decision. As noted above, the general rule of patentable subject matter is expansive and any determination of failure to claim statutory subject matter must find its support in the case law. In the final office action the Examiner concedes that the claimed invention does not lack utility under section 101 of the patent statute. Therefore, the rejection is either based on the printed matter exception or on the algorithm exception to the rule of patentability. As noted above, storage of data (which is by its nature descriptive) is a very important aspect of the information technology arts. That fact was recognized in In re Lowry when the Federal Circuit laid to rest the doctrine of printed matter as applied to data stored in computer readable media. The Federal Circuit's decision in In re Lowry requires reversal of the Examiner's determination of failure to claim statutory subject matter.

To the extent that the Examiner's section 101 rejection relied on the mathematical

algorithm doctrine it must also be reversed. In AT&T, supra, the Federal Circuit said that any step-by-step process involves an algorithm in the broad sense of the word. The AT&T court thus said: "Since the process of manipulation of numbers is a fundamental part of computer technology, we have had to reexamine the rules that govern the patentability of such technology. The sea-changes in both law and technology stand as a testament to the ability of law to adapt to new and innovative concepts, while remaining true to basic principles." AT&T, 172 F.3d at 1356. Thus the AT&T court limited the "Algorithm" doctrine to apply only in cases of purely "abstract" algorithms. See AT&T at 1357. In AT&T, the Federal Circuit also said that the algorithm must be applied in a useful way and found a practical result in the claimed methods in the addition of certain descriptive information called a PIC (or primary interchange carrier) to certain other information used in switching telephone calls. The information in the claims at issue in the instant case also has a useful result – the storage for retrieval of information from a computer memory responsive to a search for certain criteria. The retrieved information is useful for, among other purposes, determining properties of molecules.

The Examiner's argument that the information stored according to the claims at issue is merely descriptive is also not directed at an exception to patentability that is recognized by the courts. For example, if the "merely descriptive" argument were applied to the field of photography such an argument would preclude the patentability of cameras because cameras take light, one form of information that represents an object, and record the information on film. The information is merely descriptive of the subject of the photograph. The application of the Examiner's reasoning to the clearly patentable area of photography illustrates the point that the claimed invention which is analogous to other forms of data storage should not be precluded from patentability. Therefore, this rejection should be reversed.

The Federal Circuit rejected an argument similar to the Examiner's Arrhythmia

Research Technology, Inc. v. Corazonix Corp., 22 USPQ2d 1033 (Fed. Cir. 1992), where processing information describing a patient's heartbeat was held to be statutory subject matter. The court there said that the claims at issue did not preempt all uses of the algorithm, Arrhythmia at 1060. Similarly, in the instant case the claims do not preempt all uses of any algorithms; rather they are limited to storage and retrieval in a computer memory. Therefore, appellants request reversal of the rejection of the claims at issue under 35 U.S.C. §101.

II. Rejection of Claims 1, 4, and 5 Under 35 U.S.C. §§102(a) and (e)(2)

Claims 1, 4, and 5 were rejected under 35 U.S.C. §§102 (a) and (e)(2) as being anticipated by Platt (U.S. Pat. No. 5,784,294). This rejection should be reversed because the Examiner has not shown that claims 1, 4, and 5 are anticipated by Platt. Nowhere does Platt teach or disclose *any* of the elements of claim 1. Platt relates to a storage device that performs a plurality of functions that produce a result that can be an input to the method of claim 1 of the instant application but which does not anticipate the claims at issue. Platt does not disclose the required mapping, generation of a key, or storing the entry as required by claim 1. Claims 4 and 5 are dependent on claim 1 and are not anticipated by Platt for at least the same reasons.

In the final office action, the Examiner contends that Platt discloses at Fig. 9 “storing an entry comprising a molecular descriptor with a key to access it.” Fig. 9 of Platt is a flow chart illustrating use of descriptors. It does not relate to a mapping of descriptor vectors (as claimed) at all. The key generated according to claim 1 corresponds to the mapping. The Examiner has not shown how Fig. 9 of Platt performs a mapping or any of the claimed steps.

The Examiner further says “Platt et al. teaches storing said first and second descriptors of each molecule in said series of molecules in a database for subsequent processing to thereby identify correspondence between molecules in said series of molecules (Claim

34, Lines 39-42).” That statement does not describe any of the claimed steps. The claimed step of “storing” relates to the entry defined in the first step. The section of Pratt cited has nothing to do with such an entry and hence cannot correspond to the claimed storing step, or any other claimed step.

The Examiner also argues that the “key” is inherent. Again, the claimed “key” corresponds to the claimed mapping and the Examiner has not shown anything in Platt corresponding to such a mapping. Instead, the Examiner argues that Platt has some criteria for selecting molecules of the training set to be placed in a table and that this corresponds to “applying the mapping.” The Examiner does not show how the placement of molecules in a table relates even remotely to applying a mapping to a descriptor vector as claimed and has fallen far short of the exact relationship that anticipation requires.

Finally, the Examiner has erred as a matter of law by arguing that a type of data structure allegedly disclosed in Platt “is consistent with” the limitation of “key indexes to entry for retrieval thereof.” The legal test for anticipation is whether every element of a claim is found in an item of prior art, and not whether a structure is consistent with a claimed method. Therefore, appellants request reversal of the rejections under Section 102.

III. Rejections Of Claims 1, 4-15, And 31-35 Under 35 U.S.C. § 112, First Paragraph

The final Office Action rejected claims 1, 4-15, and 31-35 under 35 U.S.C. § 112, first paragraph, as failing to comply with the written description requirement. The test for compliance with the written description requirement is whether the claimed subject matter was described in the original specification as to convey to one skilled in the art that the inventor possessed the claimed invention. Vas-Cath Inc., v. Mahurkar, 935 F.2d 1555, 19 USPQ2d 111 (Fed. Cir. 1991). The examiner (or the Board, if the Board

is the first body to raise a particular ground for rejection) "bears the initial burden . . . of presenting a prima facie case of unpatentability." In re Oetiker, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). Insofar as the written description requirement is concerned, that burden is discharged by "presenting evidence or reasons why persons skilled in the art would not recognize in the disclosure a description of the invention defined by the claims." In re Wertheim, 541 F.2d 257, 263, 191 USPQ 90, 97 (CCPA 1976). Thus, the burden placed on the examiner varies, depending upon what the applicant claims. The specification contains a description of the claimed invention, albeit not in *ipsis verbis* (in the identical words), then the examiner or Board, in order to meet the burden of proof, must provide reasons why one of ordinary skill in the art would not consider the description sufficient. Id. at 264, 191 USPQ at 98.

The Examiner erred in rejecting claim 1 as not complying with the written description requirement for using the term "based upon preselected criteria."

The amendment of November 13, 2002, amended claim 1 so that the step of applying a mapping to the descriptor vector is based on pre-selected criteria. Support for the amendment is inherent from the discussion in page 40 of the specification that the application of the mapping is based on pre-determined criteria. Note that the discussion (page 40) of the "association criteria" is defined in the prior training phase which occurs before the applying a mapping step and thus clearly the association criteria were "pre-determined."

The Examiner erred in rejecting claim 1 as not complying with the written description requirement for using the term "such that the key indexes the entry for retrieval thereof."

Claim 1 was amended to state that "the key indexes the entry for retrieval thereof." It is inherent in the claimed invention that the key indexes the entry for retrieval of the entry. Why else would a key corresponding to a mapping be used? In any case, the gist of the written description requirement is to prevent an applicant from adding claims to subject

matter that the inventor did not possess at the time of filing. Vas-Cath Inc., v. Mahurkar, supra. As appellants noted in the amendment of November 13, 2002, the amendment was not made to define additional subject matter but to make clear what was already implicit. Appellants again pose the question – why would information be stored if not for retrieval thereof?

The Examiner erred in rejecting claim 4 as not complying with the written description requirement for using the term "computed from a convolution with a probe"

The Examiner contends at page 6 of the final rejection, that specific to claim 4, lines 2-3 the introduction of "computed from a convolution with a probe" is considered new matter because it has not been found in the application as filed. Appellant disagrees. The computation of a convolution is found at page 29, line 15 – page 30, line 5. Line 5 is the definition of a convolution. In any case, a smearing function is used in a convolution. See Protein Crystallography Course, page 10 (attached as Exhibit A). With respect to using a probe function that is also a well known concept. This methodology applied to electromagnetic fields was presented in J. D. Jackson's "Classical Electrodynamics, 2nd Ed." Wiley, NY 1975. The idea is that you have a piece of equipment (a probe) that you can put in the material and measure the field with. The probe has some geometric extent, and adds together parts of the field in the vicinity of the point identifying the location of the field. That type of addition is called a convolution. The instrument is called a probe. The response of the probe is a probe function.

The Examiner erred in rejecting claim 6 as not complying with the written description requirement for using the terms " a ratio of variations ... discriminant criterion function" and "a criterion function ...utilizes the first data and the second data"

The Examiner raised at page 6 of the final rejection of June 29, 2005, that claim 6, lines 5-6 recites "a ratio of variations ... discriminant criterion function" and lines 9-10 recite

"a criterion function ...utilizes the first data and the second data" which the examiner did not find in the specification. The examiner noted that pages 10-11 describe "F-distributed ratios ... representing discrimination between groups of items" which the examiner contends to be different. First the language "identifying a set of component vectors that maximizes a ratio of variations between groups to the variations within the groups along the component vectors as a discriminant criterion function" is found at page 23, line 15 of the original specification. The computation of a component vector is shown at that line.

Support for the language "generating a criterion function for subsets of the component vectors, wherein the criterion function utilizes the first data and the second data" is found at page 9, line 8 – page 9, line 4; page 22, lines 21- 22; and page 23, lines 16-22; and the step 217 of FIG. 2.

The Examiner has not shown any reason why the language added in the amendment would not be supported by the specification and in fact appellants contend that the amendment was not made for purposes of patentability, so the invention defined by the claims both before and after the amendment are the same and hence was clearly in the possession of the inventor at the time of the filing of the application. Therefore, appellants request reversal of this rejection.

IV. Rejection Of Claims 8, 9, And 12-14 Under 35 U.S.C. §112, Second Paragraph

The Examiner erred in rejecting claims 8, 9, and 12-14 under 35 U.S.C. §112, second paragraph, as being indefinite. The Examiner rejected claim 8 on grounds that " W^T " is not defined. The claims must be read in view of what those skilled in the art would understand. Phillips v. AWH Corp., 415 F.3d 1303, 1313 (Fed. Cir. 2005) (en banc). The Examiner raised at page 6 of the Office Action of June 29, 2005, that claim 8, lines 4-5, the limitation of "T indicates a transpose ... representing co-variance" has not been

found in the specification. It does not matter that the specification has not defined these terms because those terms are well known in mathematics to have the meaning shown in the claim. The specification at page 12, line 14 states that " W^T is the transpose of the unit vector w ." A copy of a page from J. Truss, *Discrete Mathematics for Computer Scientists, 2d Ed.*, Addison-Westley (1999) is attached (Exhibit B) showing the usage of the T as a well known symbol for a transpose.

The Examiner also rejected claim 8 on grounds that ε_b and ε_w are unclear because it is unclear what parameters are represented by these symbols. The term ε_b is a first data representing covariance and that is found at page 6 of the original specification. ε_w is another data representing covariance. The claim states this. The computation of the covariances ("T indicates a transpose... ..representing covariance") is spelled out on page 8, lines 3 and 6 in the original specification. This section describes in detail the specific covariances in question, what T means, etc. It even calls them covariances. That disclosure is sufficient to define the terms. Claim 9 was rejected for being dependent on claim 8. Therefore if the rejection of claim 8 is overcome so is the rejection of claim 9.

Claim 8 was further rejected for using the term "based upon" for characterizing C. The words "based" and "upon" have ordinary meanings that are applicable in the context used in the claims. The term "based" is defined as "to find a basis for." *The American Heritage® Dictionary of the English Language, Fourth Edition Copyright © 2000 Houghton Mifflin Company*. The basis for C is the degrees of freedom in ε_b and ε_w . Moreover, claim 12 specifies that C is a constant. The Examiner has not shown why such ordinary meaning renders the claims indefinite.

The rejection of claim 11 in the office action of June 10, 2004 was dropped in the final office action of June 29, 2005 and therefore the issues of the original rejection are conceded.

Claim 12 was rejected on the basis that "represents" is indefinite because the claim does not define the metes and bounds of that term. The word "represents" must be given its plain meaning as understood by those skilled in the art. The American Heritage Dictionary, Second College Edition defines "represent" as "to be the equivalent of." The American Heritage Dictionary, Second College Edition, Houghton Mifflin (1985). That is exactly the context in which it appears in claim 12 therefore the term is not indefinite.

V. Objection to Amendment of September 14, 2004.

The originally filed application contained an incorporation by reference reading as follows:

"The present application claims priority to Provisional U.S. Patent Application No. 06/079,196, and is related to U.S. Patent Applications (Attorney Docket YOR999-149) and (Attorney Docket YOR999-150), herein incorporated by reference in their entirety." It is clear that the incorporation by reference referred to "their" in the plural and that means that it was not just to the provisional application.

CONCLUSION

In view of the foregoing, it is respectfully submitted that the application and the claims are in condition for allowance. Reversals of the final rejection, and allowance of the claims as amended, are requested.

Respectfully submitted,

Date: July 17, 2006

By: Michael J. Buchenhorner
Michael J. Buchenhorner
Registration No. 33,162
Attorney for Appellants

CLAIMS APPENDIX

1. In a data processing system wherein descriptor vectors associated with a plurality of regions of molecules are stored in a database, a method for generating and storing data characterizing at least one region of said plurality of regions, the method comprising the steps of:

generating an entry comprising

- i) an identifier that identifies said at least one region, and
- ii) data characterizing a set of axes derived from a property distribution of said at least one region;
- iii) a set of axes derived from property distribution information of the at least one region, the set of axes characterizing the at least one region;

generating at least one descriptor vector for the at least one region;

applying a mapping to the descriptor vector associated with said at least one region based on preselected criteria; and

storing said entry in a memory, wherein said key is associated with said entry such that the key indexes the entry for retrieval thereof.

Claims 2 and 3 have been cancelled.

4. The method of claim 1, wherein said property distribution of said at least one region is computed from a convolution with a probe function to a property field.

5. The method of claim 1, wherein said plurality of descriptor vectors are classified into groups, and wherein said mapping step maps said descriptor vectors to a space discriminating between said groups of descriptor vectors.

6. The method of claim 5, wherein said mapping is derived from the steps of:
generating first data representing differences between said groups of descriptor vectors;

generating second data representing variations within said groups of descriptor vectors;

identifying a set of component vectors that maximizes a ratio of variations between groups to the variations within the groups along the component vectors as a discriminant criterion function;

generating a criterion function for subsets of the component vectors, wherein the criterion function utilizes the first data and the second data;

for each particular subset of component vectors, calculating a probability value for the criterion functions associated with the particular subset;

selecting a probability value from probability values for said subsets of component vectors based upon a predetermined criterion;

identifying the subset of component vectors associated with the selected probability value; and

generating a mapping to a space corresponding to the subset of component vectors associated with the selected probability value, and storing the mapping for subsequent processing.

7. The method of claim 6, wherein said first data comprises a matrix ϵ_b representing covariance between said groups of descriptor vectors, and said second data comprises a matrix ϵ_w representing covariance within said groups of descriptor vectors.

8. The method of claim 7, wherein said criterion function has the general form:

$$J(\hat{w}) = C \left(\frac{\hat{w}^T \epsilon_b \hat{w}}{\hat{w}^T \epsilon_w \hat{w}} \right)$$

where \hat{w} is some vector, T indicates a transpose, ϵ_b is a first data representing covariance, ϵ_w is a second data representing covariance and C is a constant based upon degrees of freedom in ϵ_b and ϵ_w .

9. The method of claim 8, wherein C is determined as follows:

$$C = \frac{1/\text{degrees of freedom in } \epsilon_b}{1/\text{degrees of freedom in } \epsilon_w} = \frac{1/(N-1)}{1/(\sum n_i - N)}$$

where N represents the number of groups of descriptor vectors, n_i represents the number of regions, and $\sum n_i$ represents the sum of n_i for the N groups.

10. The method of claim 7, wherein the step of identifying a set of component vectors that maximizes an F distributed criterion function comprises the substeps of:

determining a set of (eigenvalue, eigenvector) pairs for the matrix ϵ_w

determining said set of component vectors based upon said set of (eigenvalue, eigenvector) pairs for the matrix ϵ_w .

11. The method of claim 10, wherein said statistic for a given subset of component vectors is based upon value of said criterion function for said subset of component vectors.

12. The method of claim 11, wherein said statistic for a given subset of component vectors has the following form:

$$\Psi_s = C \left(\frac{1}{L_s} \right) \sum J_k$$

where J_k represents the value of the criterion function at a component vector in the given subset,

C is a constant,

L_s represents the number of J_k values in the given subset of component vectors, and the \sum operation sums over the L_s J_k values in the given subset of component vectors.

13. The method of claim 12, wherein said a probability value for a particular F-distributed statistic represents a probability value that the particular F-distributed statistic could have been larger by chance.

14. The method of claim 13, wherein said probability value selected from probability values for said subsets of component vectors is a minimum probability value of said probability values for said subsets of component vectors.

15. The method of claim 6, wherein said mapping for said at least one descriptor vector performs a loop over each component vector belonging to the subset of component vectors associated with the selected probability; wherein, in each iteration of said loop, dot product of said descriptor vector with a transpose of a unit vector for the given component vector is added to a running sum.

Claims 16-30 have been cancelled.

31. The method of claim 1, wherein the at least one descriptor vector is invariant to rotation and translation of the at least one region.

32. The method of claim 31, wherein the set of axes is derived from principal axes of second moments of a region of the property distribution information.

33. The method of claim 6, wherein the probability value is obtained by treating the ratio as an F-distributed statistic.

34. The method of claim 6, wherein the probability value is obtained by any one of cross-validation, jack-knife and bootstrap estimations.

35. The method of claim 6, wherein application in constructing the discriminant criterion function includes boosting and bagging techniques.

EVIDENCE APPENDIX



Protein Crystallography Course

Course Homepage Basic: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) Advanced: [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) [21](#) [22](#)

Search

The convolution theorem and its applications

Outline

- What is a convolution?
 - Convolutions are commutative
- The convolution theorem
 - Proof of the convolution theorem
- The correlation theorem
 - What is a correlation function?
 - Parseval's theorem
- Important convolutions
 - Convolution with a Gaussian
 - Convolution with a delta function
- Applications of the convolution theorem
 - Atomic scattering factors
 - B-factors
 - Diffraction from a lattice
 - Diffraction from a crystal
 - Resolution truncation
 - Density modification
- Applications of the correlation theorem
 - The Patterson function
 - The phased translation function

What is a convolution?

One of the most important concepts in Fourier theory, and in crystallography, is that of a convolution. Convolutions arise in many guises, as will be shown below. Because of a mathematical property of the Fourier transform, referred to as the convolution theorem, it is convenient to carry out calculations involving convolutions.

But first we should define what a convolution is. Understanding the concept of a convolution operation is more important than understanding a proof of the convolution theorem, but it may be more difficult!

Mathematically, a convolution is defined as the integral over all space of one function at x times another function at $u-x$. The integration is taken over the variable x (which may be a 1D or 3D variable),

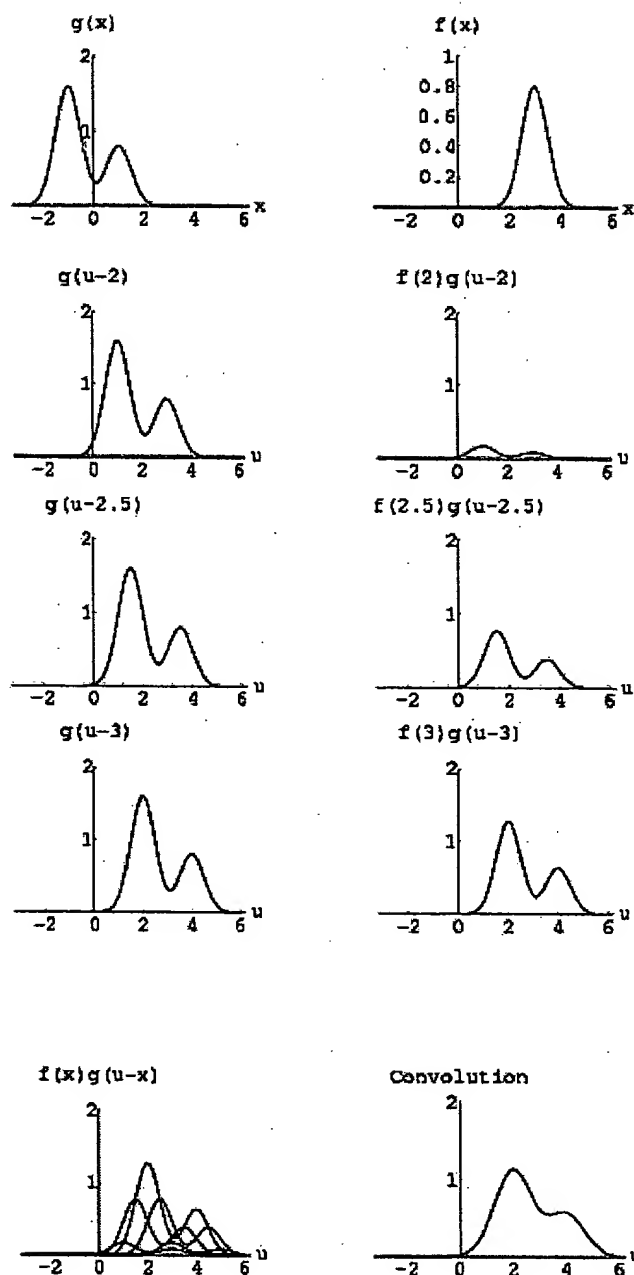
EXHIBIT

A

typically from minus infinity to infinity over all the dimensions. So the convolution is a function of a new variable u , as shown in the following equations. The cross in a circle is used to indicate the convolution operation.

$$\begin{aligned} C(u) &= f(\mathbf{x}) \otimes g(\mathbf{x}) = \int_{space} f(\mathbf{x}) g(u - \mathbf{x}) d\mathbf{x} \\ &= g(\mathbf{x}) \otimes f(\mathbf{x}) = \int_{space} g(\mathbf{x}) f(u - \mathbf{x}) d\mathbf{x} \end{aligned}$$

Note that it doesn't matter which function you take first, *i.e.* the convolution operation is commutative. We'll prove that below, but you should think about this in terms of the illustration below. This illustration shows how you can think about the convolution, as giving a weighted sum of shifted copies of one function: the weights are given by the function value of the second function at the shift vector. The top pair of graphs shows the original functions. The next three pairs of graphs show (on the left) the function g shifted by various values of x and, on the right, that shifted function g multiplied by f at the value of x .



The bottom pair of graphs shows, on the left, the superposition of several weighted and shifted copies of g and, on the right, the integral (*i.e.* the sum of all the weighted, shifted copies of g). You can see that the biggest contribution comes from the copy shifted by 3, *i.e.* the position of the peak of f .

If one of the functions is unimodal (has one peak), as in this illustration, the other function will be shifted by a vector equivalent to the position of the peak, and smeared out by an amount that depends on how sharp the peak is. But alternatively we could switch the roles of the two functions, and we would see that the bimodal function g has doubled the peaks of the unimodal function f .

Convolutions are commutative

$$f(x) \otimes g(x) = g(x) \otimes f(x)$$

We stated that the convolution integral is commutative. Here, in case you are interested, is a quick proof of that. First, we start with the convolution integral written one way. For convenience we will deal with the 1D case, but the 3D case is exactly analogous.

$$C(u) = \int_{-\infty}^{\infty} f(x)g(u-x)dx$$

Now we substitute variables, replacing $u-x$ with a new x' .

$$x' = u - x, \quad dx' = -dx$$

$$C(u) = - \int_{\infty}^{-\infty} f(u-x')g(x')dx'$$

Note that, because the sign of the variable of integration changed, we have to change the signs of the limits of integration. Because these limits are infinite, the shift of the origin (by the vector u) doesn't change the magnitude of the limits.

Now we reverse the order of the limits, which changes the sign of the equation, and swap the order of the functions g and f .

$$C(u) = \int_{-\infty}^{\infty} g(x')f(u-x')dx'$$

It doesn't matter whether we call the variable of integration x' or x , so we put back x , to get the result we wanted to prove.

$$C(u) = \int_{-\infty}^{\infty} g(x)f(u-x)dx$$

The convolution theorem

Because there will be so many Fourier transforms in the rest of this presentation, it is useful to introduce a shorthand notation. T will be used to indicate a forward Fourier transform, and its inverse to indicate the inverse Fourier transform.

$$T(f(\mathbf{r})) = \int_{\text{space}} f(\mathbf{r}) \exp(2\pi i \mathbf{s} \cdot \mathbf{r}) d\mathbf{r}$$

There are two ways of expressing the convolution theorem:

- The Fourier transform of a convolution is the product of the Fourier transforms.
- The Fourier transform of a product is the convolution of the Fourier transforms.

$$T(f \otimes g) = T(f) T(g)$$

$$T(fg) = T(f) \otimes T(g)$$

The convolution theorem is useful, in part, because it gives us a way to simplify many calculations. Convolutions can be very difficult to calculate directly, but are often much easier to calculate using Fourier transforms and multiplication.

Proof of the convolution theorem

To prove the convolution theorem, in one of its statements, we start by taking the Fourier transform of a convolution. What we want to show is that this is equivalent to the product of the two individual Fourier transforms. Note, in the equation below, that the convolution integral is taken over the variable x to give a function of u . The Fourier transform then involves an integral over the variable u .

$$\begin{aligned} T(f(x) \otimes g(x)) &= T\left(\int_{-\infty}^{\infty} f(x) g(u-x) dx\right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) g(u-x) dx \exp(2\pi i s u) du \end{aligned}$$

Now we substitute a new variable w for $u-x$. As above, the infinite integration limits don't change. Then we expand the exponential of a sum into the product of exponentials and rearrange to bring together expressions in x and expressions in w .

$$\begin{aligned} T(f(x) \otimes g(x)) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) g(w) \exp[2\pi i s(x+w)] dx dw \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) \exp(2\pi i s x) g(w) \exp(2\pi i s w) dx dw \end{aligned}$$

Expressions in x can be taken out of the integral over w so that we have two separate integrals, one over x with no terms containing w and one over w with no terms containing x .

$$T(f(x) \otimes g(x)) = \int_{-\infty}^{\infty} f(x) \exp(2\pi i s x) dx \int_{-\infty}^{\infty} g(w) \exp(2\pi i s w) dw$$

The variables of integration can have any names we please, so we can now replace w with x , and we have the result we wanted to prove.

$$\begin{aligned} T(f(x) \otimes g(x)) &= \int_{-\infty}^{\infty} f(x) \exp(2\pi i s x) dx \int_{-\infty}^{\infty} g(x) \exp(2\pi i s x) dx \\ &= T(f(x)) T(g(x)) \end{aligned}$$

If you look through the derivation above, you will see that we could have used a minus sign in the exponential when taking the original Fourier transform, and then the two Fourier transforms at the end would also contain minus signs in the exponentials. In other words, the convolution theorem applies to both the forward and reverse Fourier transforms. This is not surprising, since the two directions of Fourier transform are essentially identical.

Proof of second statement of convolution theorem

To prove the second statement of the convolution theorem, we start with the version we have already proved, *i.e.* that the Fourier transform of a convolution is the product of the individual Fourier transforms.

$$T(f \otimes g) = T(f) T(g)$$

First we'll define some shorthand, where capital letters indicate the Fourier transform mates of lower case letters.

$$\begin{aligned} F &= T(f) & f &= T^{-1}(F) \\ G &= T(g) & g &= T^{-1}(G) \end{aligned}$$

We use these relationships to recast the statement above in terms of the Fourier transform mates of the original functions. Then we take an inverse Fourier transform on each side of the equation, to get (essentially) the second statement of the convolution theorem. The only difference is that it is expressed in terms of the inverse Fourier transform.

$$T[T^{-1}(F) \otimes T^{-1}(G)] = F G$$

$$T^{-1}(F) \otimes T^{-1}(G) = T^{-1}(F G)$$

But, as we noted above, we could have proved the convolution theorem for the inverse transform in the same way, so we can reexpress this result in terms of the forward transform.

$$T(f g) = T(f) \otimes T(g)$$

The correlation theorem

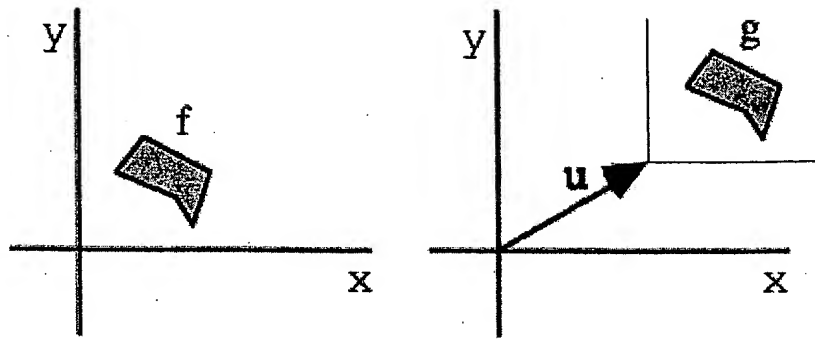
The correlation theorem is closely related to the convolution theorem, and it also turns out to be useful in many computations. The correlation theorem is a result that applies to the correlation function, which is an integral that has a definition reminiscent of the convolution integral.

What is a correlation function?

In a correlation integral, instead of taking the value of one function at $u-x$, you take the value of that function at $x-u$. Equivalently, you take the value of the other function at $x+u$. This is shown in the following equation, along with the variable substitution that allows the two expressions to be interconverted.

$$\begin{aligned} f \circ g &= \int_{-\infty}^{\infty} f(x) g(x+u) dx; \text{ substitute } x' = x+u \\ &= \int_{-\infty}^{\infty} f(x'-u) g(x') dx' \\ &= \int_{-\infty}^{\infty} f(x-u) g(x) dx \end{aligned}$$

The figure below illustrates why the correlation function has the name that it does. If the two functions f and g contain similar features, but at a different position, the correlation function will have a large value at a vector corresponding to the shift in the position of the feature.



The correlation theorem can be stated in words as follows: the Fourier transform of a correlation integral is equal to the product of the complex conjugate of the Fourier transform of the first function and the Fourier transform of the second function. The only difference with the convolution theorem is in the presence of a complex conjugate, which reverses the phase and corresponds to the inversion of the argument $u-x$.

$$T(f \circ g) = F^* G$$

Parseval's theorem

Important convolutions

Convolution with a Gaussian

First we need to define a Gaussian function. We will stick, for the moment, to 1D Gaussians. The Gaussian function is the familiar bell-shaped curve, with a peak position (r_0) and standard deviation.

$$p(r) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r-r_0)^2}{2\sigma^2}\right)$$

We won't derive the Fourier transform of a Gaussian, but it is given by the following equation.

$$T(p(r)) = \exp(2\pi i r_0 s) \exp(-2\pi^2 \sigma^2 s^2)$$

Note that the Fourier transform of a Gaussian is another Gaussian (although lacking the normalisation constant). There is a phase term, corresponding to the position of the center of the Gaussian, and then the negative squared term in an exponential. Also notice that the standard deviation has moved from the denominator to the numerator. This means that, as a Gaussian in real space gets broader, the corresponding Gaussian in reciprocal space gets narrower, and vice versa. This makes sense, if you think about it: as the Gaussian in real space gets broader, contributions from points within that Gaussian start

to interfere with each other at lower and lower resolutions.

Convolution with a Gaussian will shift the origin of the function to the position of the peak of the Gaussian, and the function will be smeared out, as illustrated above.

Convolution with a delta function

Delta functions have a special role in Fourier theory, so it's worth spending some time getting acquainted with them. A delta function is defined as being zero everywhere but for a single point, where it has a weight of unity.

$$\delta(\mathbf{r} - \mathbf{r}_0) = \begin{matrix} \text{weight of 1 at } \mathbf{r} = \mathbf{r}_0 \\ 0 \text{ elsewhere} \end{matrix}$$

What it means to say that it has a weight of unity is that the integral of the delta function over all space is 1. The delta function is given an argument of $\mathbf{r} - \mathbf{r}_0$ so that it can be defined as having its non-zero point at the origin. When \mathbf{r} is equal to \mathbf{r}_0 , the argument of the delta function is zero.

$$\int_{\text{space}} \delta(\mathbf{r} - \mathbf{r}_0) d\mathbf{r} = 1$$

A more general property of the delta function is that the integral of a delta function times some other function is equal to the value of that other function at the position of the delta function.

$$\int_{\text{space}} f(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_0) d\mathbf{r} = f(\mathbf{r}_0)$$

How can a single point, with no width, breadth or depth, have a weight of one? The value of the delta function at that point must be a special kind of infinity, and this means that it has to be defined as a limit. There are a number of ways to define a delta function. One of them is to define it as an infinitely sharp Gaussian. The integral over all space of a Gaussian is 1, which satisfies one of the properties required for the delta function, and if we take the limit of a Gaussian as the standard deviation tends to zero, it satisfies the other properties. The following equation defines a 3D delta function as the limit of an isotropic 3D Gaussian.

$$\delta(\mathbf{r} - \mathbf{r}_0) = \lim_{\sigma^2 \rightarrow 0} \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_0|^2}{2\sigma^2}\right)$$

With this definition of the delta function, we can use the Fourier transform of a Gaussian to determine the Fourier transform of a delta function. As the standard deviation of a Gaussian tends to zero, its Fourier transform tends to have a constant magnitude of 1. All that is left is the phase shift term.

$$T[\delta(\mathbf{r} - \mathbf{r}_0)] = \exp(2\pi i \mathbf{s} \cdot \mathbf{r}_0)$$

So we see that the Fourier transform of a delta function is just a phase term. Think about the picture we had of an electron at a point; it contributed just a phase term, with unit weight, to the diffraction pattern. So now we see that we can consider an electron at a point to be a delta function of electron density.

Finally we can consider the meaning of the convolution of a function with a delta function. If we write down the equation for this convolution, and bear in mind the property of integrals involving the delta function, we see that convolution with a delta function simply shifts the origin of a function.

$$\int_{\text{space}} \delta(\mathbf{r} - \mathbf{r}_0) f(\mathbf{u} - \mathbf{r}) d\mathbf{r} = f(\mathbf{u} - \mathbf{r}_0)$$

$$\delta(\mathbf{r} - \mathbf{r}_0) \otimes f(\mathbf{r}) = f(\mathbf{r} - \mathbf{r}_0)$$

Applications of the convolution theorem

Atomic scattering factors

We have essentially seen this before. We can tabulate atomic scattering factors by working out the diffraction pattern of different atoms placed at the origin. Then we can apply a phase shift to place the density at the position of the atom. Our new interpretation of this is that we are convoluting the atomic density distribution with a delta function at the position of the atom.

B-factors

We can think of thermal motion as smearing out the position of an atom, *i.e.* convoluting its density by some smearing function. The B-factors (or atomic displacement parameters, to be precise) correspond to a Gaussian smearing function. At resolutions typical of protein data, we are justified only in using a single parameter for thermal motion, which means that we assume the motion is isotropic, or equivalent in all directions. (In crystals that diffract to atomic resolution, more complicated models of thermal motion can be constructed, but we won't deal with them here.)

Above, we worked out the Fourier transform of a 1D Gaussian.

$$T(p(r)) = \exp(2\pi i r_0 s) \exp(-2\pi^2 \sigma^2 s^2)$$

In fact, all that matters is the displacement of the atom in the direction parallel to the diffraction vector, so this equation is suitable for a 3D Gaussian. All we have to remember is that the term corresponding to the standard deviation refers only to the direction parallel to the diffraction vector. Since we are dealing with the isotropic case, the standard deviation (or atomic displacement) is equal in all directions.

The B-factor is used in an equation in terms of $\sin\theta/\lambda$ instead of the diffraction vector, because all that matters is the magnitude of the diffraction vector. We replace the variance (standard deviation squared) by the mean-square displacement of the atom in any particular direction. The B-factor can be defined in terms of the resulting equation.

$$|s| = \frac{1}{d} = 2 \sin \theta / \lambda$$

$$T(p(\mathbf{r})) = \exp\left(-8\pi^2 \langle u_x^2 \rangle \sin^2 \theta / \lambda^2\right)$$

$$B = 8\pi^2 \langle u_x^2 \rangle$$

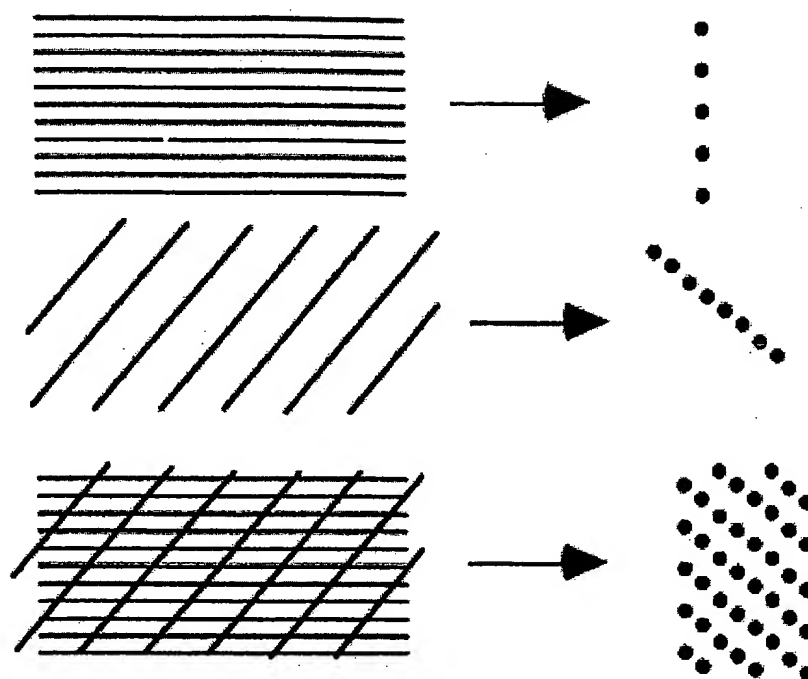
Note that there is a common source of misunderstanding here. The mean-square atomic displacement refers to displacement in any particular direction. This will be equal along orthogonal x, y and z axes. But often we think of the mean-square displacement as a radial measure, *i.e.* total distance from the mean position. The mean-square radial displacement will be the sum of the mean-square displacements along x, y and z; if these are equal it will be three times the mean-square displacement in any single direction. So the B-factor has a slightly different interpretation in terms of radial displacements.

$$\langle u_r^2 \rangle = \langle u_x^2 \rangle + \langle u_y^2 \rangle + \langle u_z^2 \rangle = 3 \langle u_x^2 \rangle$$

$$B = \frac{8\pi^2}{3} \langle u_r^2 \rangle$$

Diffraction from a lattice

The convolution theorem can be used to explain why diffraction from a lattice gives



If you've gotten this far, I'm sorry that I ran out of time to complete this document before giving the lecture! The rest will be filled in at some point after all the lectures (and associated web pages) are finished.

Diffraction from a crystal

Resolution truncation

Density modification

Solvent flattening

Sayre's equation

Applications of the correlation theorem

The Patterson function

The phased translation function

© 1999-2005 Randy J Read, University of Cambridge. All rights reserved.

Last updated: 7 June, 2005

The difference between this and Dijkstra's algorithm is that in choosing which vertex and edge to adjoin, we minimize a_{ij} , not $d_j (= d_i + a_{ij})$. We can implement Prim's algorithm to take about n^2 steps in a similar way, by using two arrays, one to tell us the current minimum weight on an edge from i to a point of S (only relevant for $i \notin S$), and another to tell us which previous vertex i is joined to; the value of the latter is fixed once $i \in S$.

Example 8.10

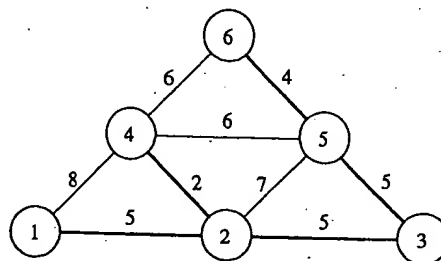


Figure 8.34

We apply Prim's algorithm to the edge-weighted graph shown in Figure 8.34. The edges given by the algorithm are shown in bold.

S	d_1	d_2	d_3	d_4	d_5	d_6	previous vertex to					
							1	2	3	4	5	6
1		5	∞	8	∞	∞	1			1		
1,2			5	2	7	∞	1	2	2	2		
1,2,4			5		6	6	1	2	2	4	4	
1,2,3,4					5	6	1	2	2	3	4	
1,2,3,4,5						4	1	2	2	3	5	
1,2,3,4,5,6							1	2	2	3	5	

Exercises 8.3

- Show that a relation having adjacency matrix A is
 - reflexive if and only if $I + A = A$
 - symmetric if and only if $A^T = A$ (where A^T , the transpose of A , is obtained by interchanging its rows and columns)
 - transitive if and only if $A + A^2 = A$ where 'addition' is taken as max on $\{0, 1\}$.
- Find the transitive closures of the relations R_1 and R_2 on $\{0, 1, 2, \dots, 20\}$ given by

$$(x, y) \in R_1 \text{ if } y = x + 3$$

$$(x, y) \in R_2 \text{ if } y = x + 3 \pmod{21}$$
 Is either of R_1 and R_2 an equivalence relation? What about $TC(R_1)$ or $TC(R_2)$?

EXHIBIT

B

BEST AVAILABLE COPY

RELATED PROCEEDINGS APPENDIX

NONE

3909410_v2

RELATED PROCEEDINGS APPENDIX

NONE

3909410_v2